

K-Means

Instances (Örnekler)	X	Y
1	1.0	1.5
2	1.0	4.5
3	2.0	1.5
4	2.0	3.5
5	3.0	2.5
6	5.0	6.0

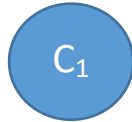
Örnek : k =2 seçilsin ve 2 iterasyon yapılsın.

Adım 1 : Ortalamayı hesaplamak için rastgele iki nokta (k=2) belirlenir. Bu noktalar,

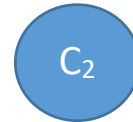
C_1 : (1.0 , 1.5) ve C_2 : (2.0 , 1.5) olsun (veri seti üzerinden rastgele k tane nokta seçildi).

Adım 2 : Öklit, Manhattan veya Minkowski uzaklık formüllerinden biriyle her bir satır için hesaplama yapılır. Burada öklit uzaklık formülünü kullanacağız.

C1		C2
$d(C_1, I_1) = \sqrt{(1.0 - 1.0)^2 + (1.5 - 1.5)^2} = 0$	<	$d(C_2, I_1) = \sqrt{(2.0 - 1.0)^2 + (1.5 - 1.5)^2} = 1$
$d(C_1, I_2) = \sqrt{(1.0 - 1.0)^2 + (1.5 - 4.5)^2} = 3$	<	$d(C_2, I_2) = \sqrt{(2.0 - 1.0)^2 + (1.5 - 4.5)^2} = 3.1$
$d(C_1, I_3) = \sqrt{(1.0 - 2.0)^2 + (1.5 - 1.5)^2} = 2.24$	>	$d(C_2, I_3) = \sqrt{(2.0 - 2.0)^2 + (1.5 - 1.5)^2} = 0$
$d(C_1, I_4) = \sqrt{(1.0 - 2.0)^2 + (1.5 - 3.5)^2} = 2.24$	>	$d(C_2, I_4) = \sqrt{(2.0 - 2.0)^2 + (1.5 - 3.5)^2} = 2$
$d(C_1, I_5) = \sqrt{(1.0 - 3.0)^2 + (1.5 - 2.5)^2} = 2.24$	>	$d(C_2, I_5) = \sqrt{(2.0 - 3.0)^2 + (1.5 - 2.5)^2} = 1.4$
$d(C_1, I_6) = \sqrt{(1.0 - 5.0)^2 + (1.5 - 6.0)^2} = 6.02$	>	$d(C_2, I_6) = \sqrt{(2.0 - 5.0)^2 + (1.5 - 6.0)^2} = 5.4$



I1 , I2



I3 , I4 , I5 , I6

Görüldüğü gibi öklit uzaklık sonucu hangi kümenin küçükse o Instance (Örnek) o kümeye aittir.

Adım 3 : Kümelenen Instance'ların nitelikleri ile ortalama hesaplanır.

$$C_1 = (\frac{X_1 + X_2}{2}, \frac{Y_1 + Y_2}{2})$$

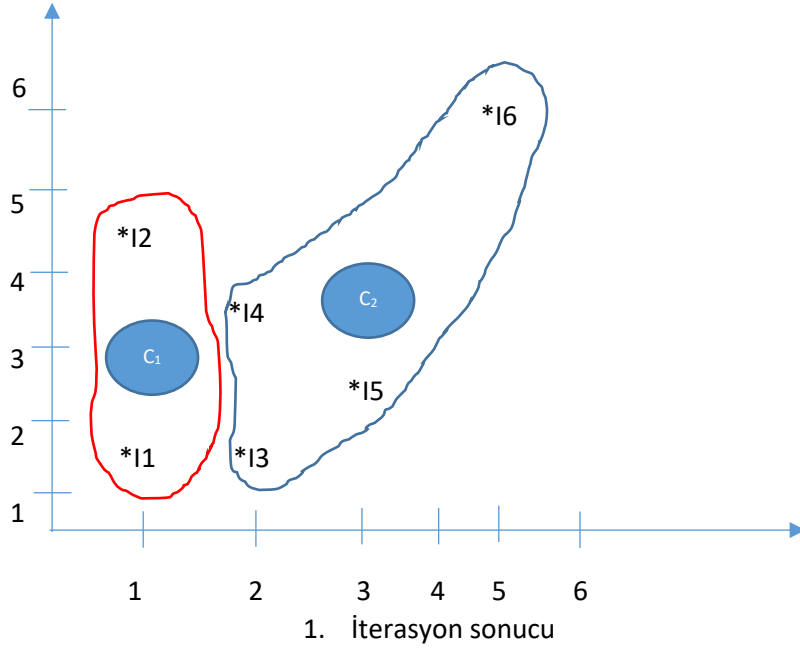
$$C_1 = (\frac{1.0 + 1.0}{2}, \frac{1.5 + 4.5}{2})$$

$$C_1 = (1, 3)$$

$$C_2 = (\frac{X_3 + X_4 + X_5 + X_6}{4}, \frac{Y_3 + Y_4 + Y_5 + Y_6}{4})$$

$$C_2 = (\frac{2.0 + 2.0 + 3.0 + 5.0}{4}, \frac{1.5 + 3.5 + 2.5 + 6.0}{4})$$

$$C_2 = (3, 3.38)$$



Adım 4 : İlk iterasyon tamamlandı. Şimdi 2. İterasyon için yeni c_1 ve c_2 noktaları kullanılarak tekrar öklit uzaklık formülü yardımıyla kümeleme işlemi yapılır ve ortalamalar bulunur.

$$d(c_1, I1) = \sqrt{(1.0 - 1.0)^2 + (3.0 - 1.5)^2} = 1.5 < d(c_2, I1) = \sqrt{(3.0 - 1.0)^2 + (3.38 - 1.5)^2} = 2.75$$

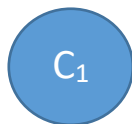
$$d(c_1, I2) = \sqrt{(1.0 - 1.0)^2 + (3.0 - 4.5)^2} = 1.5 < d(c_2, I2) = \sqrt{(3.0 - 1.0)^2 + (3.38 - 4.5)^2} = 2.25$$

$$d(c_1, I3) = \sqrt{(1.0 - 2.0)^2 + (3.0 - 1.5)^2} = 1.8 < d(c_2, I3) = \sqrt{(3.0 - 2.0)^2 + (3.38 - 1.5)^2} = 2.14$$

$$d(c_1, I4) = \sqrt{(1.0 - 2.0)^2 + (3.0 - 3.5)^2} = 1.11 > d(c_2, I4) = \sqrt{(3.0 - 2.0)^2 + (3.38 - 3.5)^2} = 1.0$$

$$d(c_1, I5) = \sqrt{(1.0 - 3.0)^2 + (3.0 - 2.5)^2} = 2.06 > d(c_2, I5) = \sqrt{(3.0 - 3.0)^2 + (3.38 - 2.5)^2} = 0.90$$

$$d(c_1, I6) = \sqrt{(1.0 - 5.0)^2 + (3.0 - 5.5)^2} = 5.0 > d(c_2, I6) = \sqrt{(3.0 - 5.0)^2 + (3.38 - 5.5)^2} = 3.20$$



I1 , I2 , I3



I4 , I5 , I6

Görüldüğü gibi öklit uzaklık sonucu hangi kümenin küçükse o Instance (Örnek) o kümeye aittir.

$$C_1 = (X_1 + X_2 + X_3 / 3, Y_1 + Y_2 + Y_3 / 3)$$

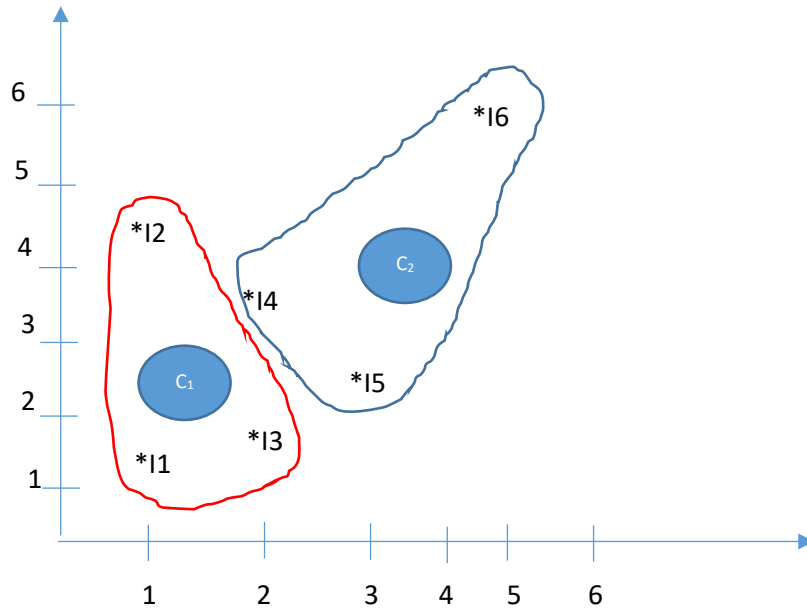
$$C_1 = (1.0 + 1.0 + 2.0 / 3, 1.5 + 4.5 + 1.5 / 3)$$

$$C_1 = (1.33, 2.5)$$

$$C_2 = (X_4 + X_5 + X_6 / 3, Y_4 + Y_5 + Y_6 / 3)$$

$$C_2 = (2.0 + 3.0 + 5.0 / 3, 3.5 + 2.5 + 6.0 / 3)$$

$$C_2 = (3.33, 4)$$



Hakan Cem Gerçek

gmail: hakancg95gmail.com

instagram : hkn.cem

Twitter: eightjune95